NATIONAL QUALITY FORUM

National Voluntary Consensus Standards for Patient Outcomes
Summary of the Cancer Technical Advisory Panel Conference Call
March 29, 2010

TAP members: Lee Newcomer, MD, MHA; Susan Beck, PhD, APRN, FAAN; Christopher Friese, RN, PhD, AOCN; Patricia Haugen; David Penson, MD, MPH; Bonnie Teschendorf, PhD, MHA; Joe B. Putnam, MD, FACS; Susan Goodin, PharmD, FCCP, BCOP

NQF staff: Heidi Bossley, Sarah Fanta, Hawa Camara

Measure Steward Representatives: Bruce Hall (ACS); Jennifer Beaumont (FACIT); Ronald Walters (MD Anderson Cancer Center); Cary Kaufman (National Consortium of Breast Cancer Centers); Lawrence Bassett (UCLA School of Medicine)

Audience:

Dr. Newcomer began the call with welcome and introductions by the Technical Advisory Panel (TAP) members. TAP members were asked to disclose any conflict with the measures being discussed.

Ms. Bossley, NQF senior director, performance measures, provided an introductory slide presentation that described

- NQF and its activities;
- The HHS funded patient outcomes project;
- The role of the TAP;
- NQF's standard measure evaluation criteria; and
- Identifying gaps in outcomes measures.

Dr. Newcomer led TAP members through discussion of the sub-criteria for the five submitted measures. Measure developers were present and responded to questions from TAP members. The rating and issues discussed are summarized in the tables that follow.

OT2-010-09: Imaging timeliness of care-time between diagnostic mammogram and needle/core biopsy

IMPORTANCE TO MEASURE AND REPORT		
1a Impact	Overall Ranking: Minimally Completely: 0 Partially: 2 Minimally: 5 Did Not	Members were concerned that this measure on timeliness did not have a demonstrated link to an outcome – there was no evidence provided that it impacted survival/mortality or addressed disparities in care (e.g., delays). In addition, it was unclear how this information would help inform the public. The measure developer clarified that the primary issue that the measure is trying to address is the anxiety that

	Meet: 1	women experience and also be patient-centered. Another potential concern is how the measure truly addresses care coordination since often the diagnosis is made at one institution but the biopsy occurs at another. It was unclear how the measure addresses coordination and to whom attribution/accountability is applied. The measure developer clarified that the measure is intended to enhance collaboration but the measure is attributed to the biopsy entity. The TAP stated that additional research on how these types of measures relate to timeliness of care and reduced timeframes between diagnosis and treatment impacted outcomes – particularly reducing the anxiety of a patient – would be needed to enable the measure to be considered further. It was determined that the measure did not pass the importance criteria; thus, it was not discussed further.
1b Gap	Overall Ranking: Minimally Completely: 1 Partially: 5 Minimally: 2 Did Not Meet: 0	
1c Relation to Outcomes	Overall Ranking: Did Not Meet Completely: 0 Partially: 0 Minimally: 0 Did Not Meet: 8	

OT2-011-09: Surgical timeliness of care-time between needle biopsy and initial breast cancer surgery

IMPORTANCE TO MEASURE AND REPORT		
1a Impact	Overall Ranking: Minimally Completely: 0 Partially: 2 Minimally: 6 Did Not Meet: 0	The TAP raised similar concerns as was discussed for Measure #OT2-010-09. An additional area of concern was raised that the measure as written may have an unintended negative consequence of not encouraging patient preferences and time to make an informed decision due to the time window included in the measure. It is critical that delays do not impact outcomes but it must also not discourage time for an adequate assessment and for patients to determine the best treatment option. Many other factors that may be appropriate care including patients who determine that they would like a second opinion or coordination with plastic surgery for reconstruction may also impact the ability to comply with the measure. The measure developer emphasized that while those factors could be included as exclusions, they chose to not include
1b Gap	Overall Ranking: Minimally Completely: 0 Partially: 0	

	Minimally: 8 Did Not Meet: 0	them in the measure but rather anticipate that 100% performance may not achievable. Because this measure raised similar concerns as the previous measure, the TAP determined that the measure did not pass the Importance criteria; thus, it was not discussed further.
1c Relation to Outcomes	Overall Ranking: Did Not Meet Completely: 0 Partially: 0 Minimally: 0 Did Not Meet: 8	

OT2-016-09: Functional assessment of cancer therapy-lung (FACT-L)
OT2-017-09: Functional assessment of breast therapy-breast (FACT-B)
OT2-019-09: Functional assessment of cancer therapy-general version (FACT-G)

Note: Given the similarity with these three measures, the following summary discusses all three.

IMPORTANCE TO MEASURE AND REPORT		
1a Impact	Overall Ranking: Partially Completely: 3 Partially: 5 Minimally: 0 Did Not Meet: 0	All agreed that these surveys were well described, validated and have been proven as excellent tools in clinical trials. They have demonstrated the ability to determine the differences in the quality of life of patients who are treated with one therapy versus another and could be used to guide patients to therapeutic interventions that improve quality of life. Some concerns were raised by members. It is not clear how the tool translates into the care of individual patients at the point of care rather than its traditional use for populations of patients through clinical trials. It is also unclear how the tool directly relates or measures an outcome and how it would be used for public reporting. Because the tools have been primarily used in clinical trials, its application in practices is unclear. In addition, the information provided does not demonstrate that a gap in care exists. Are there variations across practices or patients?
1b Gap	Overall Ranking: Did Not Meet Completely: 0 Partially: 0 Minimally: 1 Did Not Meet: 7	
1c Relation to Outcomes	Overall Ranking:	

	Minimally Completely: 0 Partially: 2 Minimally: 5 Did Not Meet: 1	
SCIENTIFIC ACCEPTABILTY		Overall Ranking: Completely Completely: 6 Partially: 2 Minimally: 0 Did Not Meet: 0
2a Specs	Overall ranking: TBD Completely: 1 Partially: 1 Minimally: 0 Did Not Meet: 1	The tools have been well tested to address reliability and validity of measure quality of life but does it translate to a quality of care measure. Risk adjustment and exclusions should be considered if it is to be used as a quality of care measure. Stratification of the populations (e.g., treatment, stage) is critical. ***Additional Cancer TAP Comments following call:*** *Fundamental question: should QOL measures be used for quality of care measures? I do not believe we have enough developmental science for endorsement. If used as QOC measure, recommend careful attention to risk adjustment and exclusion criteria.*
2b Reliability	Overall Ranking: Completely Completely: 2 Partially: 1 Minimally: 0 Did Not Meet: 0	
2c Validity	Overall Ranking: Completely Completely: 2 Partially: 1 Minimally: 0 Did Not Meet: 0	
2d Exclusions	Overall Ranking: TBD Completely: 1 Partially: 0	

	Minimally: 1 Did Not Meet: 1	
2e Risk Adjustment	Overall Ranking: TBD Completely: 1 Partially: 0 Minimally: 1 Did Not Meet: 1	
2f Meaningful Differences	Overall Ranking: Completely Completely: 2 Partially: 0 Minimally: 0 Did Not Meet: 1	
2g Comparability	Overall Ranking: Completely Completely: 2 Partially: 0 Minimally: 0 Did Not Meet: 1	
2h Disparities	Overall Ranking: Completely Completely: 2 Partially: 0 Minimally: 0 Did Not Meet: 1	
USEABILITY		Overall Ranking: Minimally Completely: 0 Partially: 1 Minimally: 5 Did Not Meet: 2
3a Distinctive	Overall Ranking: Partially Completely: 0 Partially: 2 Minimally: 1 Did Not Meet: 0	It is unclear how this tool will inform patients when publicly reported. It is valuable to use as a static tool using a patient as a control or in a patient population that is well aligned. How the data would be used for public reporting is unclear and stratification will be necessary to generate useful comparisons.

3b Harmonization	Overall Ranking: TBD Completely: 0 Partially: 1 Minimally: 1 Did Not Meet: 1	
3c Added Value	Overall Ranking: Partially Completely: 0 Partially: 2 Minimally: 1 Did Not Meet: 0	
FEASIBILITY		Overall ranking: Minimally Completely: 0 Partially: 1 Minimally: 6 Did Not Meet: 0
4a Data a Byproduct of Care	Overall Ranking: Did Not Meet Completely: 0 Partially: 0 Minimally: 0 Did Not Meet: 2	These tools have been widely used but not routinely used in clinical practice, which represents a challenge to feasibility. Particularly, small practices and indigent communities may find it difficult to implement them unless they are electronically available. The availability of the tool in multiple languages was viewed as a strength. ***Additional Cancer TAP Comments following call:*** *Exclusions must be addressed for qoc perspective*
4b Electronic	Overall Ranking: Did Not Meet Completely: 0 Partially: 0 Minimally: 0 Did Not Meet: 2	
4c Exclusions	Overall Ranking: TBD Completely: 1 Partially: 1 Minimally: 0 Did Not Meet:0 Not Applicable: 1	
4d	Overall	

Inaccuracies/Errors	Ranking: TBD Completely: 1 Partially: 1 Minimally: 0 Did Not Meet:1	
4e Implementation	Overall Ranking: TBD Completely: 0 Partially: 1 Minimally: 1 Did Not Meet:1	

OT2-002-09: Risk adjusted colorectal surgery (Outcome Measure)

IMPORTANCE TO MEASURE AND REPORT		
1a Impact	Overall Ranking: Completely Completely: 5 Partially: 3 Minimally: 0 Did Not Meet: 0	This measure is meaningful as mortality and severe morbidity are important. The information provided and intent of the measure clearly meets the subcriteria for importance.
1b Gap	Overall Ranking: Completely Completely: 6 Partially: 2 Minimally: 0 Did Not Meet: 0	
1c Relation to Outcomes	Overall Ranking: Partially Completely: 2 Partially: 6 Minimally: 0 Did Not Meet: 0	
SCIENTIFIC ACCEPTABILTY		Overall ranking: Partially

		Completely: 0 Partially: 8 Minimally: 0 Did Not Meet: 0
2a Specs	Overall Ranking: Partially Completely: 1 Partially: 2 Minimally: 0 Did Not Meet: 0	Information on reliability and validity of the measure is provided. The ability to compare across hospitals is a real strength. One concern raised was on how smaller hospitals will be able to publicly report on this measure given the need for a sufficient number of cases. The measure developer clarified that approximately 65 cases were needed each year, which means that the measure may only apply to 40 to 45 percent of all hospitals but would cover 85 percent of all colorectal surgery. As included in the forms, reliability was found to be moderate. The measure developer clarified that while the findings were found to be moderate, the information is more than typically provided and meet acceptable standards proposed in the literature. A member also questioned whether the measure has been validated outside of NSQIP but the developer supports that the measure can be implemented in other programs and by other organizations. It is estimated that it would require about a 20th of a full-time employee to abstract the data needed for the measure. ***additional cancer TAP commentsl:*** *Strong methodology. Lack of validation outside NSQIP data platform. Reliance on x numbers of patients per hospital still yields poor reliability. Learner curve for data abstractors outside of NSQIP platform. These issues must be addressed before endorsement can be considered.*
2b Reliability	Overall Ranking: TBD Completely: 1 Partially: 1 Minimally: 1 Did Not Meet: 0	
2c Validity	Overall Ranking: TBD Completely: 1 Partially: 0 Minimally: 1 Did Not Meet: 1	
2d Exclusions	Overall Ranking: Partially Completely: 0 Partially: 2 Minimally: 1 Did Not Meet: 0	
2e Risk Adjustment	Overall Ranking: Completely Completely: 3 Partially: 0 Minimally: 0 Did Not Meet:	

	0	
2f Meaningful Differences	Overall Ranking: TBD Completely: 1 Partially: 1 Minimally: 1 Did Not Meet: 0	
2g Comparability	Overall Ranking: Minimally Completely: 0 Partially: 0 Minimally: 3 Did Not Meet: 0	
2h Disparities	Overall Ranking: Partially Completely: 0 Partially: 2 Minimally: 0 Did Not Meet: 1	
USEABILITY	Overall Ranking: Partially Completely: 0 Partially: 6 Minimally: 1 Did Not Meet: 1	
3a Distinctive	Overall Ranking: Minimally Completely: 0 Partially: 1 Minimally: 2 Did Not Meet: 0	Members of the TAP were unsure of whether the public would understand the composite nature of the measure. In addition, the performance data provided appeared to show that the improvement curve is relatively flat. The developer clarified that those who have implemented the measure have demonstrated improvement over time. Consumer understanding of the odds ratio is sometimes difficult and may impact its usability for patients.
3b harmonization	Overall Ranking: TBD	***Additional cancer TAP comments::***

	Completely: 1 Partially: 0 Minimally: 1 Did Not Meet: 0	*Does not specifically address why 30-day mortality or failure to rescue are not suitable for measurement (already NQF endorsed). Innovations appear to be risk adjustment. However, this measure is more difficult to collect, measure, and report. I contend that NSQIP mortality rates are not uniformly improving.*
3c Added value	Overall Ranking: Minimally Completely: 0 Partially: 1 Minimally: 2 Did Not Meet: 0	
FEASIBILITY	Overall ranking: Minimally Completely: 0 Partially: 2 Minimally: 6 Did Not Meet: 0	
4a Data a byproduct of care	Overall Ranking: Did Not Meet Completely: 0 Partially: 0 Minimally: 1 Did Not Meet: 2	One of the key concerns for the TAP was feasibility and the associated costs to implement the measure. All agreed that the developer has demonstrated how the measure works but it is not yet clear whether it can also be implemented outside of a hospital that currently participates in NSQIP. The data required must be generated by abstraction and is not readily available through electronic data sources. The measure developer clarified that it is a parsimonious algorithm that hospitals would apply and other organizations would be able to implement this measure without participation in NSQIP. ***Additional cancer TAP comments:*** *Risk of mis-measurement largely unknown as collected currently in NSQIP. Validation study in non-NSQIP hospitals on feasibility, usability, reliability, and validity recommended prior to final endorsement.*
4b Electronic	Overall Ranking: Minimally Completely: 0 Partially: 1 Minimally: 2 Did Not Meet: 0	
4c Exclusions	Overall Ranking: Partially Completely: 0 Partially: 2 Minimally: 0 Did Not Meet:	

	1	
4d Inaccuracies/Errors	Overall Ranking: Minimally Completely: 1 Partially: 0 Minimally: 2 Did Not Meet: 0	
4e Implementation	Overall Ranking: Minimally Completely: 0 Partially: 0 Minimally: 2 Did Not Meet: 1	